

DESCRIPTIONMULTI PERSPECTIVE VIDEO CAPTURE SYSTEMTECHNICAL FIELD

5           The present invention relates to a system for acquiring video information and a storage medium and, more particularly, to a multi perspective video capture system for capturing and storing picture information afforded from the multiple viewpoints, a storage medium  
10 for a program that controls the multi perspective video capture system, and a storage medium for storing video information.

BACKGROUND ART

15           In a variety of fields such as sport in addition to manufacturing and medicine, a physical body in the real world is captured by a processor and a variety of processes may be attempted on a processor. For example, information on the movement of a person or thing and the  
20 shape of the physical body is captured and used in the analysis of the movement of the person or thing and in the formation of imaginary spaces, and so forth.

          However, because operations are performed in a variety of environments, the person or physical body that  
25 is to be actually evaluated is not necessarily in a place that is suitable for capturing information. Further, in order to capture the phenomenon in which the real world

is made with a processor as is, it is necessary to not generate obstacles to the operation and not take time for target objects such as people and objects and the peripheral environment thereof.

5           Conventionally, a procedure known as motion capture is known as the procedure for capturing an object in such a real world on a computer. This motion capture simulates the movement of a moving body such as a person. As a motion capture device, Japanese Patent Kokai Publication No. 10   2000-321044 (paragraph numbers 0002 to 0005), for example, is known. Japanese Patent Kokai Publication No. 2000-321044 mentions, as motion capture systems, optical, mechanical, and magnetic systems that are known as representative examples of motion capture, for example, 15   and in the motion capture of an optical system, a marker is attached in the location in which the movement of the body of an actor is to be measured and the movement of each portion is measured from the position of the marker by imaging the marker by means of a camera. In mechanical 20   motion capture, an angle detector and pressure-sensitive device are attached to the body of the actor and the movement of the actor is detected by detecting the bend angle of the joints. In magnetic motion capture, a magnetic sensor is attached to each part of the actor's 25   own body, the actor is moved in an artificially generated magnetic field and the actor's movement is detected by deriving the absolute position in which the magnetic

sensor exists by detecting the density and angle of the lines of magnetic force by means of a magnetic sensor.

#### DISCLOSURE OF THE INVENTION

5           In the case of conventionally known motion capture, the attachment of special markers in positions in which the body of the test subject is determined in an optical system, the placement of a camera in the periphery of the target object on the basis of homogeneous light, the  
10           placement of the target object in an artificially generated magnetic field in a magnetic system, the attachment of an angle detector and pressure sensitive device to the body of the test subject in a mechanical system, and the fact that calibration (correction), which  
15           performs correction of the actual position and pixel positions in the camera image, takes time, and so forth, necessitate a special environment, and there is the problem that the burden on the test subject and party performing the measurement is great.

20           In addition, in conventional motion capture, positional information of only representative points determined for the target object is measured, and movement is detected on that basis. Picture information for the target object is not included. Although conventional  
25           motion capture of an optical system comprises a camera, this camera acquires position information on markers that are attached in representative positions from an image

of a target object such as a test subject, the image data of the target object is discarded, and the original movement of the target object is not captured. As a result, the movement of the target object that is obtained in  
5 conventional motion capture is represented in a wire-frame form, for example, and there is a problem that the original movement of the target object cannot be reproduced.

Furthermore, in a conventional system, a high-cost  
10 camera is required in order to capture an image of the target object highly accurately and a more expensive camera is required in order to capture an image of a wide area in particular.

In order to acquire the position and attitude of  
15 the target object by using images that are picked up by a video camera, it is necessary to analyze the position and attitude of the target object that is photographed over individual frames for a row of images (frames). The analytical accuracy generally increases as larger  
20 photographs of the subject are taken. The reason is that the shift in the position of the real world of the subject is reflected as a shift in the position on the frame (pixel position) as the proportion of the subject with respect to the viewing angle increases.

25 One method for increasing the accuracy is a method for increasing the accuracy of the pixels on the frame. However, this method limits the performance of the pickup

element of the video camera and is confronted by the problem that the data amount of the image transmission increases excessively and is therefore not practical. Therefore, in order to capture a large subject, the cameraman may  
5 move (pan, tilt) the viewing field of the camera or zoom in. In addition, the camera itself may also be moved in accordance with the movement of the subject.

However, when the camera parameters such as pan, tilt, zoom, and the movement of the camera itself are  
10 changed during photography, there is the problem that analysis of the position and attitude is impossible. In a normal analysis method, data that is known as the camera parameters such as the spatial position, line of sight, breadth of field (found from the focal length) of the  
15 camera are initially captured and a calculation formula (calibration formula) for combining the camera parameters and the results of the image analysis on individual frames (position on the subject) is created to calculate the subject's position in the real world. In addition, a  
20 space position can be estimated by performing this calculation on frame data of two or more video cameras. In such calculation of the subject's position, when the camera parameters change during photography, it is not possible to accurately calibrate the image data.

25 Therefore, the present invention resolves the above conventional problem and an object thereof is to acquire the actual movement including a picture image of the target

object independently of the measurement environment. A further object is to acquire a wide-range picture highly accurately without using a highly expensive camera.

The present invention reduces the burden on a target  
5 object such as a test subject by acquiring multi  
perspective video images data by photographing the target  
object by means of a plurality of cameras and acquires  
the actual movement including a picture of the target  
object independently of the measurement environment by  
10 acquiring camera parameters such as the attitude and zoom  
of the camera along with picture data.

The present invention acquires video image data by  
synchronizing a plurality of cameras during photographing  
by the cameras and at the same time acquires camera  
15 parameters for each frame in sync with the video image  
data, rather than simply acquiring video image data and  
camera parameters, and therefore is capable of acquiring  
the actual movement of the target object independently  
of the measurement environment and of acquiring the  
20 movement of the picture itself of the target object rather  
than movement of only representative points.

The present invention comprises the respective  
aspects of a multi perspective video capture system (multi  
perspective video image system) for acquiring video  
25 information of the target object from multi perspective,  
a storage medium for a program that causes a computer  
to execute control to acquire video information of the

target object from multi perspective, and a storage medium for storing video information of the target object acquired from multi perspective.

5 A first aspect of the multi perspective video capture system (multi perspective video image system) of the present invention is a video capture system that acquires video information on a target object from multi perspective, wherein mutually association information is added to video image data that is acquired from a  
10 plurality of cameras that operate in sync with one another and to the camera parameters of each camera to the data is outputted. The outputted video image data and camera parameters can be stored and picture data and camera parameters are stored for each frame.

15 A second aspect of the multi perspective video capture system of the present invention is a video capture system that acquires video information of the target object from multi perspective that is constituted comprising a plurality of cameras for acquiring moving  
20 images, detector for acquiring the camera parameters of each camera; synchronizer for acquiring moving images by synchronizing a plurality of cameras; data appending device that make associations between the video data of each camera and between the video image data and camera  
25 parameters.

Video image data is acquired by synchronizing a plurality of cameras by means of the synchronizer means

and respective video image data acquired by each camera are synchronized by the data appending device and the video image data and camera parameters are synchronized. As a result, the video image data and camera parameters of a plurality of cameras of the same time can be found.

Furthermore, the second aspect further comprises video image data storage for storing video image data rendered by adding association information for each frame and camera parameter storage for storing camera parameters rendered by adding association information. According to this aspect, video image data and camera parameters including mutually association information can be stored. Further, for the video image data storage and camera parameter storage, different storage or the same storage can be assumed. Further, when the same storage are used, video image data and camera parameters can each be stored in different regions or can be stored in the same region.

In the above aspect, it can be assumed that the association information is the frame count of video image data that is acquired by one camera of a plurality of cameras. By referencing the frame count, the association between the respective frames of the video image data that is acquired from a plurality of cameras is known and, in addition to being able to process picture data at the same time in sync, camera parameter data that corresponds with the video image data of the same time

can be found and processed in sync.

The camera parameters contain camera attitude information of camera pan and tilt and zoom information. Pan is the oscillation angle in the lateral direction of the camera, for example, and tilt is the oscillation angle in the vertical direction of the camera, for example, where pan and tilt are attitude information relating to the imaging directions in which the camera performs imaging. Further, the zoom information is the focal position of the camera, for example, and is information relating to the viewing field range that is captured on the imaging screen of the camera. The attitude information of the camera makes it possible to know the pickup range in which the camera performs imaging in accordance with zoom information.

The present invention comprises, as camera parameters, zoom information in addition to the camera attitude information of pan and tilt and is therefore able to obtain both an increase in the resolution of the video image data and an enlargement of the acquisition range.

In addition, the multi perspective video capture system of the present invention can also include two-dimensional or three-dimensional position information for the camera as the camera parameters. On account of including the position information, even in a case where the camera itself has moved in a space, the

spatial relationship between the picture data of each camera can be grasped and picture information can be acquired over a wide range with a small number of cameras. In addition, image information can be acquired while  
5 tracking a moving target object.

Further, in addition to the above camera parameters, the data that is stored for each frame can also be data of every kind such as measurement data and measured measurement data can be stored in sync with picture data  
10 and camera parameters.

An aspect of the program storage medium of the present invention is a storage medium for a program that causes a computer to execute control to acquire video information of a target object from multi perspective, comprising  
15 first program encoder that sequentially add a synchronization common frame count to video image data of each frame acquired from a plurality of cameras; and second program encoder that sequentially add a frame count corresponding to the video image data to the camera  
20 parameters of each camera.

The first program encoder include the storage in first storage of picture data to which a frame count has been added and the second program encoder include the storage in second storage of count parameters to which  
25 a frame count has been added. This program controls processing executed by the data appending device.

Furthermore, the camera parameters include the

camera attitude information of camera pan and tilt and zoom information. Further, the camera parameters may include camera two-dimensional or three-dimensional position information. In addition, for example, a  
5 variety of information on the photographic environment and periphery such as sound information, temperature, and level of humidity may be associated and stored with video image.

As a result of a constitution in which other  
10 information is associated and stored with video image data in addition to the camera parameters, a sensor for measuring the body temperature, the outside air temperature, and a variety of gases, for example, is provided on the clothes and measurement data that is formed  
15 by these sensors in addition to the video image data imaged by the camera is captured and then associated and stored with video image data, whereby video image data and measurement data at the same time can be easily analyzed.

Furthermore, the present invention is able to  
20 correct a shift in the camera parameters that results when the camera pans and tilts. This correction comprises the steps of acquiring an image in a plurality of rotational positions by panning and/or tilting a camera; finding correspondence between the focal position of the camera  
25 and the center position of the axis of rotation from the image; acquiring the camera parameters of the camera; and correcting the camera parameters on the basis of the

correspondence.

An aspect of the storage medium of the video information of the present invention is a storage medium for storing video information of the target object that is acquired from multi perspective that stores first video information rendered by sequentially adding a synchronization common frame count to the video image data of the respective frames that is acquired from a plurality of cameras and second video information produced as a result of sequentially adding the frame count corresponding with video image data to the camera parameters of each camera. The camera parameters may include camera attitude information of camera pan and tilt and zoom information and may include camera two-dimensional or three-dimensional position information. Further, a variety of information that is associated with the video image data may be included.

It is possible to acquire video information without adding restrictive conditions such as the limited space of a studio or the like in order to render the measurement environment homogeneous light and facilitate correction.

The video information acquired by the present invention can be applied to the analysis of the movement and attitude and so forth of the target object.

25

#### BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a constitutional view to illustrate an

overview of the multi perspective video capture system of the present invention;

Fig. 2 shows an example of a constitution in which the multi perspective video capture system of the present invention comprises a plurality of cameras;

Fig. 3 serves to illustrate a picture that is imaged by a camera that the multi perspective video capture system of the present invention comprises;

Fig. 4 serves to illustrate pictures that are imaged by a camera that the multi perspective video capture system of the present invention comprises;

Fig. 5 is a constitutional view that serves to illustrate the multi perspective video capture system of the present invention;

Fig. 6 shows an example of a data array on a time axis that serves to illustrate the acquisition state of video image data and camera parameters of the present invention;

Fig. 7 shows an example of video image data and camera parameters that are stored in the storage of the present invention;

Fig. 8 shows an example of the format of video image data of the present invention and camera parameter communication data;

Fig. 9 shows an example of the structure of the camera parameter communication data of the present invention;

Fig. 10 is a schematic view that serves to illustrate

the relationship between the center of revolution of the camera and the focal position of the camera;

Fig. 11 is a schematic view that serves to illustrate the relationship between the center of revolution and  
5 the focal position of the camera;

Fig. 12 is a schematic view that serves to illustrate the correction of the camera parameters in the calibration of the present invention;

Fig. 13 is a flowchart to illustrate a camera  
10 parameter correction procedure of the present invention;

Fig. 14 serves to illustrate the camera parameter correction procedure of the present invention;

Fig. 15 shows the relationship between a three-dimensional world coordinate system representing  
15 the coordinates of the real world and a camera-side two-dimensional coordinate system;

Fig. 16 serves to illustrate an example of the calculation of the center position from the focal position of the present invention;

20 Fig. 17 is an example of a reference subject of the present invention; and

Fig. 18 shows an example in which the camera of the present invention is moved three-dimensionally by means of a crane.

25

#### BEST MODE FOR CARRYING OUT THE INVENTION

An embodiment of the present invention will be

described with reference to the attached drawings.

Fig. 1 is a constitutional view to illustrate an overview of the multi perspective video capture system (multi perspective video image system) of the present invention. In Fig. 1, a multi perspective video capture system 1 comprises a plurality of cameras 2 (cameras 2A to camera 2D are shown in Fig. 1) that acquire video image data for a moving image of the target object 10; a sensor 3 for acquiring camera parameters of each camera 2 (Fig. 1 shows sensors 3A to 3D); synchronizer 4 (only a synchronization signal is shown in Fig. 1) for acquiring a moving image by synchronizing a plurality of cameras 2; and data appending device 6 that make an association between the video image data of each camera 2 and the video image data and camera parameters. Mutually association information is added to video image data that is acquired from a plurality of cameras operating in sync with each other and to the camera parameters of each camera. The resulting data is then outputted.

The association information added by the data appending device 6 can be established on the basis of the frame count extracted from the video image data of one camera, for example. The frame count can be found by a frame counter device 7 described subsequently.

Further, the multi perspective video capture system 1 can comprise video image data storage 11 for storing video image data rendered as a result of association

information being added by the data appending device 6 and camera parameter storage 12 that store camera parameters rendered as a result of association information being added by the data appending device 6.

5       The plurality of cameras 2A to 2D can be provided in an optional position in the periphery of the target object 10 and can be fixed or movable. The cameras 2A to 2D image the moving image of the target object 10 in sync by means of a synchronization signal generated by  
10   the synchronizer 4. Further, the synchronization is performed for each frame that is imaged by the camera 2 and can also be performed in predetermined frame units. As a result, video image data that is obtained from each of the cameras 2A to 2D is synchronized in frame units  
15   and becomes video image data of the same time. The video image data that is acquired by each camera 2 is collected by the data appending device 6.

      Further, sensors 3A to 3D that detect camera parameters such as zoom information such as focal length  
20   and camera attitude information such as pan and tilt for each camera are provided for each of the cameras 2A to 2D and the camera parameters detected by each sensor 3 are collected by the data collection device 5.

      The frame count used as association information  
25   makes it possible to capture video image data from one camera among a plurality of cameras 2 and to count and acquire each frame of the video image data. The acquired

frame count constitutes information to provide associations between the respective video image data by synchronizing the video image data and information for associating video image data and camera parameters.

5       The data appending device 6 add association information that is formed on the basis of the frame count to the camera parameters collected by the video image data and data collection device 5. The video image data to which the association information is added is stored  
10 in the video image data storage 11 and the camera parameters to which the association information is added is stored in the camera parameter storage 12.

Further, the multi perspective video capture system 1 of the present invention can also have a constitution  
15 that does not comprise the video image data storage 11 and camera parameter storage 12 or can have a constitution that comprises the video image storage 11 and the camera parameter storage 12.

Fig. 2 shows an example of a constitution having  
20 a plurality of cameras that the multi perspective video capture system of the present invention comprises. Further, Fig. 2 shows an example of a constitution having four cameras which are cameras 2A to 2D as the plurality of cameras but the number of cameras can be an optional  
25 number of two or more. Camera 2A will be described as a representative example.

Camera 2A comprises a camera main body 2a and a sensor

3A for forming camera parameters is provided in the camera main body 2a. The sensor 3A comprises an attitude sensor 3a, a lens sensor 3b, a sensor cable 3c, and a data relay 3d. The camera main body 2a is supported on a camera platform which rotates or turns on at least two axes such that same is free to pan (oscillation in a horizontal direction) and tilt (oscillation in a vertical direction). Further, in cases where the cameras 2 are horizontally attached to a camera platform, pan becomes oscillation in a vertical direction and tilt becomes oscillation in a horizontal direction. The camera platform can also be installed on a tripod.

The attitude sensor 3a is a sensor for detecting the direction and angle of oscillation of the camera which detects and outputs the degree of oscillation of the camera 2A as pan information and tilt information by providing the attitude sensor 3a on the camera platform. Further, the lens sensor 3b is a sensor for detecting zoom information for the camera 2A and is capable of acquiring the zoom position of the lens by detecting the focal length, for example.

The attitude sensor 3a and lens sensor 3b can be constituted by rotary encoders having a coupled axis of rotation and detect the extent of rotation in any direction (right rotation direction and left rotation direction, for example) with respect to the reference rotation position, for example, by means of the rotation direction

and rotation angle. Further, data on the rotation direction can be expressed by a positive (+) or negative (-) when the reference rotation direction is positive, for example. Further, the rotary encoder can also use  
5 the absolute angle position that is obtained by using the absolute type. Each of the camera parameters of pan, tilt and zoom that are obtained by the attitude sensor 3a and lens sensor 3b are collected by the data collecting device 5 after being collected by the data relay 3d via  
10 the sensor cable 3c.

A picture that is obtained by the cameras of the multi perspective video capture system of the present invention will be described by using Figs. 3 and 4.

Fig. 3 shows a case where a wide viewing field is  
15 photographed by adjusting the zoom of the camera and Fig. 3B shows an example of picture data. In this case, instead of a wide viewing field being obtained, the size of each image is small. As a result, a more detailed observation of a target object 10a in the target object 10, for example,  
20 is difficult.

In this state, by enlarging the target object 10a in the target object 10 by means of the zoom function of the camera, the target object 10a can be observed with a high resolution but the viewing field range in turn  
25 narrows. The multi perspective video capture system of the present invention adjusts the problem of the contrariety of the image enlargement and narrowing of

the viewing field range by using the pan and tilt camera attitude information and zoom information and secures a wider viewing field range by means of pan and tilt also in a case where an image is enlarged by means of the zoom.

5        Fig. 4 shows a state where the zoom, pan and tilt are combined. C in Fig. 4D shows an enlarged image of the target object 10a in a position in Fig. 4B. In order to widen the viewing field range narrowed by the zoom, by panning leftward as shown in Fig. 4A, for example,  
10    the leftward image shown in C-L in Fig. 4D can be acquired and, by panning rightward as shown in Fig. 4C, the rightward image shown in C-R in Fig. 4D can be acquired. Further, by tilting upward or downward, the upward and downward images shown in C-U and C-D respectively in Fig. 4D can  
15    be acquired. Further, by combining pan and tilt, the rightward upward image shown in C-R-U in Fig. 4D can be acquired.

Thereafter, a more detailed constitutional example of the multi perspective video capture system of the  
20    present invention will be described by using Figs. 5 to 9. Further, Fig. 5 is a constitutional view serving to illustrate the multi perspective video capture system. Fig. 6 shows an example of a data array on a time axis that serves to illustrate the acquisition state of picture  
25    data and camera parameters of the present invention. Fig. 7 shows an example of picture data and camera parameters that are stored in the storage of the present invention.

Fig. 8 shows an example of the format of video image data and camera parameter communication data. Fig. 9 shows an example of the structure of the camera parameter communication data.

5        In Fig. 5, the multiperspective video capture system 1 comprises a plurality of cameras 2 (Fig. 5 shows cameras 2A to 2D); sensors 3 (Fig. 5 shows sensors 3A to 3D) for acquiring camera parameters of each camera 2; synchronizer 4 (synchronizing signal generator 4a,  
10       distributor 4b) for acquiring a moving image by synchronizing the plurality of cameras 2; a data collection device 5 for collecting camera parameters from each sensor 3; data appending device 6 (communication data controller 6a and RGB superposition means 6b) that  
15       make associations between video image data of each camera 2 and between video image data and camera parameters, and a frame counter device 7 that outputs a frame count as information for making an association. The multi perspective video capture system 1 further comprises  
20       video image data storage 11 for storing video image data outputted by the data appending device 6 and camera parameter storage 12 for storing camera parameters.

      The synchronizer 4 divides the synchronization signal generated by the synchronizing signal generator  
25       4a to the respective cameras 2A to 2D by means of the distributor 4b. Each of the cameras 2A to 2D performs imaging on the basis of the synchronization signal and

performs acquisition of the video image data for each frame. In Fig. 6, Fig. 6B shows video image data that is acquired by camera A and outputs the video image data A1, A2, A3, ..., and An, in frame units in sync with the  
5 synchronization signal. Similarly, Fig. 6G displays video image data acquired by camera B and outputs the video image data B1, B2, B3, ..., and Bn in frame units in sync with the synchronization signal.

The picture data of each frame unit contains an RGB  
10 signal and SYNC signal (vertical synchronization signal), for example, and the SYNC signal counts the frames and is used in the generation of the frame count that makes associations between the frames and between the video image data and camera parameters. Further, the RGB signal  
15 may be a signal form that is either an analog signal or digital signal.

Further, the synchronization signal may be outputted in frame units or for each of a predetermined number of frames. When the synchronization signal is  
20 outputted in each of a predetermined number of frames, frame acquisition between synchronization signals is performed with the timing of each camera and frame acquisition between cameras is synchronized by means of the synchronization signal for each of a predetermined  
25 number of frames.

The data collector 5 collects camera parameters (camera pan information, tilt information, and zoom

information) that is detected by the sensors 3 (attitude sensor 3a and lens sensor 3b) provided for each camera. Further, each sensor 3 produces an output in the signal form of an encoder pulse that is outputted by a rotary encoder or the like, for example. The encoder pulse contains information on the rotation angle and rotation direction with respect to the camera platform of the pan and tilt and contains information on the movement (or rotation amount of the zoom mechanism) and direction of the zoom.

The data collector 5 captures the encoder pulse outputted by each of the sensors 3A to 3D in sync with the SYNC signal in the video image data (vertical synchronization signal) and communicates serially with the data appending device 6.

Fig. 6C shows the camera parameters of the sensor 3A that are collected by the data collector. Camera parameter PA1 is read in sync with the SYNC signal (vertical synchronization signal) of the video image data A1 and the subsequent camera parameter PA2 is read in sync with the SYNC signal (vertical synchronization signal) of the video image data A2, and reading is similarly sequentially performed in sync with the SYNC signal (vertical synchronization signal) of the respective video image data.

The SYNC signal (vertical synchronization signal) that is used as a synchronization signal when the camera

parameters are read employs video image data that is acquired from one camera among a plurality of cameras. In the example shown in Figs. 5 and 6, an example that employs the video image data of camera 2A is shown.

5           Therefore, as for the camera parameters of the sensor 3B collected by the data collector, as shown in Fig. 6H, the camera parameter PB1 is read in sync with the SYNC signal (vertical synchronization signal) of the video image data A1 and the subsequent camera parameter PB2  
10 is read in sync with the SYNC signal (vertical synchronization signal) of the video image data A2 and, similarly, reading is sequentially performed in sync with the SYNC signal (vertical synchronization signal) of the video image data An of camera 3A. As a result,  
15 synchronization of the camera parameters of the respective cameras 3A to 3D collected in the data collector 5 can be performed.

          The frame counter device 7 forms and outputs a frame count as information for making associations in each of  
20 the frame units between the video image data of each of the cameras 2A to 2D and associations in each of the frame units between the video image data and camera parameters. The frame count acquires video image data from one camera among a plurality of cameras 2, for example, and is acquired  
25 by counting each of the frames of the video image data. The capture of the video image data may employ an external signal of a synchronization signal generation device or

the like, for example, as the synchronization signal. In the example shown in Figs. 5 and 6, an example employing the video image data of the camera 2A is shown.

Fig. 6C shows a frame count that is acquired on the basis of the video image data A1 to An, .... Further, here, for the sake of expediency in the description, an example is shown in which frame count 1 is associated with the frames of the video image data A1, frame count 2 is associated with the frames of the subsequent video image data A2, and the subsequent frame counts are increased. However, the initial value of the frame count and the increased number (or reduced number) of the count can be optional. Further, the resetting of the frame counter can be performed at an optional time by operating a frame counter reset pushbutton or when the power supply is turned ON.

The data collector 5 adds the frame count to the collected count parameter and communicates the frame count to the data appending device 6.

The data appending device 6 comprise communication data controller 6a and an RGB superposition device 6b. The data appending device 6 can also be constituted by a personal computer, for example.

The communication data controller 6a receive information on the camera parameters and the frame count from the data collector 5, stores same in the camera parameter storage 12, and extracts information on the

frame count.

The RGB superposition device 6b captures video image data from each of the cameras 2A to 2D, captures the frame count from the communication data controller 6a, superposes the frame count on the RGB signal of the video image data, and stores the result in the video image data storage 11. The superposition of the frame count can be performed by rendering frame code by encoding the frame count and then adding same to the part of the scanning signal constituting the picture data that is not an obstacle to signal regeneration, for example.

Figs. 6E and 6I show a storage state in which the video image data and frame count are stored in the video image data storage. For example, where the video image data of the camera 2A is concerned, as shown in Fig. 6E, the frames of the video image data A1 are stored with frame count 1 superposed as frame code 1, the frames of the video image data A2 are stored with frame count 2 superposed as frame code 2 and, in sequence thereafter, storage is performed with the superposition of the frame codes corresponding with the video image data. Further, where the video image data of camera 2B is concerned, as shown in Fig. 6I, the frames of the video image data B1 are stored with frame count 1 superposed as frame code 1, the frames of the video image data B2 are stored with frame count 2 superposed as frame code 2 and, in sequence thereafter, storage is performed with the superposition

of the frame codes corresponding with the video image data. So too for the video image data of multiple cameras, storage is performed with the superposition of frame code corresponding with video image data. By performing  
5 storage with the superposition of frame code on the video image data, synchronization of the frame units of the respective video image data acquired by a plurality of cameras is possible.

Figs. 6F and 6J show a storage state in which the  
10 camera parameters and frame count are stored in the camera parameter storage. For example, where the camera parameters of sensor 3A are concerned, as shown in Fig. 6F, the camera parameter PA1 is stored with frame count 1 superposed as frame code 1, the frames of the camera  
15 parameter PA2 are stored with camera count 2 superposed as frame code 2 and, sequentially thereafter, storage is performed with the superposition of the frame codes corresponding with the camera parameters. Further, where the camera parameters of sensor 3B are concerned,  
20 as shown in Fig. 6J, the camera parameter PB1 is stored with frame count 1 superposed as frame code 1, the frames of the camera parameter PB2 are stored with frame count 2 superposed as frame code 2 and, in sequence thereafter, storage is performed with the superposition of the frame  
25 code corresponding with the picture camera parameters. By performing storage with the superposition of frame code on the camera parameters, synchronization of the

frame units of the video image data of a plurality of cameras and the camera parameters of a plurality of sensors is possible.

Fig. 7 shows examples of video image data that is stored by the video image data storage and examples of camera parameters that are stored by the camera parameter storage.

Fig. 7A is an example of video image data that is stored in the video image data storage and is shown for the cameras 2A to 2D. For example, the video image data of the camera 2A is stored with the video image data A1 to An and the frame codes 1 to n superposed in each of the frames.

Furthermore, Fig. 7B shows an example of camera parameters that are stored in the camera parameter storage for the sensors 3A to 3D. For example, the camera parameters of sensor 3A are stored with the camera parameters PA1 to PAn and the frame codes 1 to n superposed for each frame.

The video image data and camera parameters stored in the respective storage make it possible to extract synchronized data of the same time by using added frame codes.

An example of the data constitution of the camera parameters will be described next by using Figs. 8 and 9.

Fig. 8 shows an example of the format of the

communication data of the camera parameters. In this example, 29 bytes per packet are formed. The 0<sup>th</sup> byte HED stores header information, the first to twenty-eighth bytes A to a store data relating to the camera parameters, and the twenty-ninth byte SUM is a checksum. Data checking is executed by forming an AND from a predetermined value and the total value of the 0<sup>th</sup> byte (HED) to the twenty-seventh byte (a).

Further, Fig. 9 is an example of communication data of the camera parameters. The data of the frame count is stored as A to C, the camera parameters acquired from the first sensor are stored as D to I, the camera parameters acquired from the second sensor are stored as J to O, the camera parameters acquired from the third sensor are stored as P to U, and the camera parameters acquired from the fourth sensor are stored as V to a. Codes for the respective pan, tilt and zoom data (Pf (code for the pan information), Tf (code for the tilt information), and Zf (code for the zoom information)) are held as the camera parameters.

Camera calibration will be described next.

In order to specify a three-dimensional position, a three-dimensional position in the real world and a corresponding pixel position in a camera image must be accurately aligned. However, correct association is not possible due to a variety of factors in the real image. As a result, correction is performed by means of

calibration. As a correction procedure, a method that estimates camera parameters from a set consisting of points on an associated image and real-world three-dimensional coordinates is employed. As this method, a method known as the Tsai algorithm that finds the physical amount of the attitude and position of the camera and the focal position by also considering the distortion of the camera is known. In the case of the Tsai algorithm, a set of points on a multiple-point world coordinate system and points on image coordinates that correspond with the former points are used. As external parameters, a rotational matrix (three parameters) and parallel movement parameters (three parameters) are found and, as internal parameters, the focal length  $f$ , lens distortion  $\tau_1, \tau_2$ , scalar coefficient  $s_x$ , and image origin  $(C_x, C_y)$  are found. The rotational array, parallel movement array and focal length are for variation at the time of photography, and the camera parameters are recorded together with video image data.

Calibration is performed by photographing a reference object by means of a plurality of cameras and using a plurality of sets of points on the reference object corresponding with pixel positions on the image of the photographed reference object. The calibration procedure photographs the object whose three-dimensional position is already known, acquires camera parameters by making an association with points on the image, acquires

a target object on the image, and calculates the three-dimensional position of the target object on the basis of the camera parameters obtained by individual cameras and the position of the target object acquired  
5 on the image.

The calibration that is conventionally performed corrects the camera parameters of a fixed camera. On the other hand, in the case of the multi perspective video capture system of the present invention, pan, tilt, and  
10 zoom are performed during photography, and the camera parameters change. Thus, when the pan, tilt and zoom of the camera change, there are no new problems with the fixed camera.

Fig. 10 is a schematic view that serves to illustrate  
15 the relationship between the center of revolution of the camera and the focal position of the camera. In Fig. 10, A is the focal point of the camera, B is the center position B of the pan rotation of the camera, and C is the center position of the tilt rotation of the camera. Camera 2  
20 comprises a camera platform 13 that provides rotatable support on at least the two axes of pan and tilt, and a tripod 14 that rotatably supports the camera platform 13. Each of the center positions B, C, and D and the focal point A of the camera do not necessarily match. Hence,  
25 pan and tilt and so forth do not rotate about the focal point of the camera and instead rotate about the axis of rotation of the part that fixes the camera of the camera

platform or the like.

Fig. 11 is a schematic view that serves to illustrate the relationship between the center of revolution and the focal position of the camera. Further, the camera is described hereinbelow as being fixed accurately to the installation center position of the tripod. As shown in Fig. 11, the relationship between one point on the circumference and the center coordinate of the circle is maintained between the focal position of the camera, and the pan rotation coordinate system, and the focal position of the camera and the tilt rotation coordinate system. Fig. 11A shows the relationship between the center O of the axis of rotation and the focal position F of the camera in a case where the camera is panned, and Fig. 11B shows the relationship between the center O of the axis of rotation and the focal position F of the camera in a case where the camera is tilted.

As shown in Fig. 11, because the center O of the axis of rotation and the focal position F of the camera do not match, when rotation takes place about the center O of the axis of rotation, the focal position F of the camera is displaced in accordance with this rotation. As a result of the displacement of the focal position F, displacement is produced between a point on the photographic surface of the camera and a real three-dimensional position, an error is produced in the camera parameters thus found, and an accurate position

cannot be acquired. In order to correct the camera parameters, it is necessary to accurately determine the positional relationship of the axis of rotation and the focal point of the camera

5        Fig. 12 is a schematic view that serves to illustrate the correction of the camera parameters in the calibration of the present invention. Further, although Fig. 12 shows an example with four cameras which are the cameras 2A to 2D as the plurality of cameras, an optional number  
10 of cameras can be obtained.

      In Fig. 12, video image data is acquired from the plurality of cameras 2A to 2D and camera parameters are acquired from the sensors provided for each camera. In a picture system such as a conventional motion capture  
15 picture system, camera parameters that are acquired from each of the fixed cameras are calibrated on the basis of the positional relationship between a predetermined real position and position on an image (single dot-chain line in Fig. 12).

20        On the other hand, in the case of the multi perspective video capture system of the present invention, displacement of the camera parameters produced as a result of the camera panning, tilting, and zooming is corrected on the basis of the relationship between the camera focal  
25 position and the center position of the axis of rotation. Correction of the camera parameters is performed for each of the frames by finding the relationship between the

focal position of the camera and the center position of the axis of rotation on the basis of the camera image data, finding correspondence between the camera parameters before and after correction from this positional relationship, and converting camera parameters that are calibrated on the basis of this correspondence.

Further, the relationship between calibration and the camera focal position and center position of the axis of rotation can be acquired by imaging the reference object and is found beforehand before acquiring image data.

Thereafter, the procedure to correct the camera parameters will be described in accordance with the flowchart of Fig. 13 and the explanatory diagram of Fig. 14. Further, the number of S in Fig. 14 corresponds with the number of S in the flowchart.

In Fig. 11, if it is possible to acquire the positional coordinates of a plurality of focal points when the camera is panned (or tilted), the pan (or tilt) rotation coordinate values can be calculated and the relationship between the positional coordinates of the focal points and the pan (or tilt) rotation coordinate values can be found from the pan (or tilt) rotation coordinate values. The camera parameters acquired from the sensors are rendered with the center position of the axis of rotation serving as the reference and, therefore, camera parameters with the position of the focal point

serving as the reference can be acquired by converting the camera parameters by using this relationship.

Because the same is true for tilt, pan will be described below by way of example.

5 First, the center position of the rotation is found by means of steps S1 to S9. The pan position is determined by moving the camera in the pan direction. The pan position can be an optional position (step S1). An image is acquired in the pan position thus determined.

10 Thereupon, a reference object is used as the photographic target in order to perform calibration and correction (step S2). A plurality of images is acquired while changing the pan position. The acquired number of images can be an optional number of two or more. Fig. 14 shows

15 images 1 to 5 as the acquired image (step S3).

An image of a certain pan position is read from the acquired image (step S4) and the coordinate position ( $u, v$ ) on the camera coordinates of the reference position ( $X_w, Y_w, Z_w$ ) of the reference object is found from the image

20 thus read. Fig. 15 shows the relationship between a three-dimensional world coordinate system representing the coordinates of the real world and a two-dimensional coordinate system of a camera. In Fig. 15, the three-dimensional position  $P (X_w, Y_w, Z_w)$  in a world

25 coordinate system corresponds to  $P (u, v)$  in the camera two-dimensional coordinate system. Correspondence can be found with the reference position found on the reference

object serving as the indicator (step S5).

Which position in the real world is projected onto which pixel on the camera image can be considered according to the pinhole camera model in which all the light is collected at one point (focal point) as shown in Fig. 15, and the relationship between the three-dimensional position  $P (X_w, Y_w, Z_w)$  of the world coordinate system and  $P (u,v)$  of a two-dimensional coordinate system on a camera image can be expressed by the following matrix equation.

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ r_{21} & r_{22} & r_{23} & r_{24} \\ r_{31} & r_{32} & r_{33} & r_{34} \end{pmatrix} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix}$$

Twelve values among  $r_{11}$  to  $r_{34}$  which are unknown values in the matrix equation can be found by using at least six sets of sets of correspondence between a known point  $(X_w, Y_w, Z_w)$  and point  $(u,v)$  (step S6).

Calibration of the camera parameters is performed by correcting camera parameters by using the values  $r_{11}$  to  $r_{34}$  thus found. The camera parameters include internal variables and external variables. Internal variables include the focal length, image center, image size, and strain coefficient of the lens, for example. External variables include the rotational angles of pan and tilt and so forth and the camera position, for example. Here,

the focal position  $(x, y, z)$  of the pan position is found by the calibration (step S7).

The process of steps S4 to S7 is repeated for the image that is acquired in the process of steps S1 to S3, and the focal position in the pan position is found. Fig. 14 shows a case where the focal positions  $F1 (x1, y1, z1)$  to  $F5 (x5, y5, z5)$  are found from images 1 to 5. Further, at least three points may be found in order to calculate the center of the axis of rotation. However, the positional accuracy of the center of the axis of rotation can be raised by increasing the focal position used in the calculation (step S8).

Thereafter, the center position  $O (x0, y0, z0)$  of the pan rotation is found from the focal position thus found. Fig. 16 serves to illustrate an example of the calculation of the center position from the focal position.

Two optional points are calculated from the plurality of focal positions found and a vertical bisector is acquired as a straight line linking two points. At least two vertical bisectors are found and the center position  $O (x0, y0, z0)$  of the pan rotation is found from the point of intersection between these vertical bisectors.

Further, in cases where two or more bisectors are found, the average of the positions of the intersecting points is found and this position then constitutes the

center position  $O(x_0, y_0, z_0)$  of the pan rotation (step S9).

Because the center position  $O(x_0, y_0, z_0)$  of the pan rotation and the respective focal positions are found as a result of the above process, correspondence between the rotation angle  $\theta$  of the pan of the center position  $O$  of the pan rotation and the rotation angle  $\theta'$  of the pan of the respective focal positions can be found geometrically (step S10). The pan rotation angle is corrected on the basis of the correspondence thus found (step S11).

Although pan is taken as an example in the above description, correction can be performed in the same way for tilt.

Fig. 17 is an example of a reference object. In order to increase the accuracy of the correction, it is necessary to acquire various angles (pan angle, tilt angle) for each camera and it is desirable to acquire these angles automatically. In such automatic acquisition, in order to acquire correspondence between an actual three-dimensional position and a two-dimensional position on the photographic surface of the camera, it is necessary for a reference position to be imaged even in a case where the oscillation angle of pan and tilt is large.

For this reason, the reference object is desirably of a shape such that the reference position is reproduced

on the photographic surface even at large pan and tilt oscillation angles. The reference object 15 in Fig. 17 is an example. The reference object has an octagonal upper base and lower base, for example, the upper and lower bases being linked by side parts on two levels. The parts of each of the levels are constituted by eight square faces and the diameter of the part at which the levels adjoin one another is larger than the diameter of the upper and lower bases. As a result, each apex is a protruding state and, when the apex is taken as the reference position, position detection can be rendered straightforward. Each face may be provided with a lattice shape (checkered flag) pattern.

Further, this shape is an example and the upper and lower bases are not limited to having an octagonal shape and may instead have an optional multisided shape. In addition, the number of levels may be two or more. Even in cases where the oscillation angle of pan and tilt is increased as the number of multisided shapes and the number of levels are increased, reproduction of the reference position is straightforward on the photographic screen.

A case where the camera itself is moved in a space will be described next. By moving a camera in three-dimensions in a space, concealment of part of the reference object and photographic target can be prevented. A crane can be used as means for moving the camera in three dimensions in a space. Fig. 18 shows an example

in which the camera of the present invention is moved three-dimensionally by means of a crane.

The crane attaches an expandable rod to the head portion of a support part such as a tripod or similar and can be controlled remotely in three-dimensions while the camera always remains horizontal. Further, the pan and tilt of the camera can be controlled in the same position as the control position of the crane and the zoom of the camera can be controlled by means of manipulation via a camera control unit.

Furthermore, by providing the camera platform 17 that supports the rod with a sensor for detecting the pan angle, tilt angle and expansion, the operating parameters of the crane can be acquired and can be synchronized and stored in association with the picture data in the same way as the camera parameters.

According to the present invention, synchronized frame number data is superposed and written to the recording device as frame data (video image data) outputted by the camera at the same time as a signal (gain lock signal) for frame synchronization is sent to each camera. Similarly, pan, tilt, zoom, and position data for the camera itself are acquired from a measurement device that is mounted on the camera in accordance with a synchronization signal. Even when this camera parameter data is acquired in its entirety every time, for example, 4 byte  $\times$  6 data is acquired at a rate of

60 frames every second, meaning that this is only 14400 bits per second, which can also be transmitted by a camera by using an ordinary serial line. In addition, the camera parameter data from each camera is a data amount that  
5 can be collected adequately by using a single computer but even if around eight video cameras are used and frame numbers are added, because the data amount is extremely small at around 200 bytes at a time and 12 kilobytes per second, storage of the data amount on a recordable medium  
10 such as a disk is also straightforward. That is, even when the camera parameters are recorded separately, because the frame acquisition times and frame numbers are strictly associated, analysis is possible. In addition, according to the present invention, optional  
15 data that is acquired by another sensor such as a temperature sensor, for example, can be recorded associated with the frame acquisition time and data analysis in which correspondence with the image is defined can be performed.

20 In each of the above aspects, the camera parameters may add position information for each camera to pan information, tilt information, and zoom information of each camera. By adding the camera position information, even when the camera itself has moved, the target object  
25 and position thereof on the acquired picture data can be found and, even in a case where the target has moved over a wide range, correspondence can be implemented

without producing a range in which it is not possible to acquire video image data by means of a small number of cameras rather than installing a multiplicity of cameras.

5           Moreover, for the camera parameters, in addition to camera attitude information and zoom information, various information on the photographic environment and periphery such as sound information, temperature, and humidity may be stored associated with the video image  
10 data. For example, sensors for measuring body temperature, the outside air temperature and a variety of gases and a pressure sensor, and so forth, are provided and measurement data formed by these sensors in addition to the video image data imaged by the camera is captured,  
15 and may be stored in association with the picture data. As a result, a variety of data relating to the imaged environment such as the external environment in which people work such as the outside air temperature and atmosphere components and the internal environment such  
20 as a person's body temperature and a load such as pressure acting on each part of the person's body can be stored associated at the same time as the video image data and video image data and measurement data of the same time can be easily read and analyzed.

25           According to an aspect of the present invention, the measurement environment is homogeneous light and it is possible to acquire video information without adding

control conditions such as space that is limited to a studio in order to simplify correction.

The video information acquired by the present invention can be applied to an analysis of the movement  
5 and attitude of the target object.

As described earlier, according to the present invention, actual movement including an image of the target object can be acquired independently of the measurement environment. Further, according to the  
10 present invention, a wide-range picture can be acquired highly accurately.

#### INDUSTRIAL APPLICABILITY

The present invention can be used in the analysis  
15 of a moving body such as a person or thing and in the formation of virtual spaces and can be applied to the fields of manufacturing, medicine, and sport.